

Utilizing Weakly Trained GANs to produce 3D Models

Kelechi M. Ikegwu¹; Hsuan-Yu Chen²; Jiayi Cao³; Smit Desai⁴

Illinois Informatics Institute¹; Department of Computer Science^{2,3}; School of Information Sciences⁴

ikegwu2@illinois.edu¹; hychen2@illinois.edu²; jcao7@illinois.edu³; smitad2@illinois.edu⁴

Ranjitha Kumar

Department of Computer Science
ranjitha@illinois.edu

ABSTRACT

Generative Adversarial Networks (GAN) have become the de-facto standard for generative models. However large amounts of data are required to achieve excellent performance which makes GANs challenging to apply to sparse datasets. In order to utilize GANs with sparse datasets we propose using weakly trained GANs to produce models which we then use to augment existing data. For this study we utilize 2 classes of a popular 3D dataset called 3DShapeNet to train on. Next data augmentation is performed on the original models using our proposed method, lastly we conduct a study to assess our models on Amazon Mechanical Turk.

Author Keywords

General Adversarial Networks; 3D-Models; Data Creation; Data Augmentation

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; I.2.m. Artificial Intelligence: Miscellaneous; I.4.m. Image Processing and Computer Vision: Miscellaneous

INTRODUCTION

There are a multitude of application areas where generative models are a promising method for solving certain problems. Generative Adversarial Networks (GAN) [2] is a particular model which is becoming the standard for generative models. However these models generally require a lot of data to generate sufficient results.

As an attempt to solve this problem we propose using the latent spaces from weakly trained GANs with sparse datasets to augment existing data. In the subsequent sections we frame our work with respect to research that has been conducted

with GANs, we discuss about our methodology for performing transformations to our data, producing models, and performing an evaluation of our produced models.

RELATED WORKS

The foundation of this research is based on prior work of several key papers. We use a rich 3D model repository called ShapeNet [1] which contains a multitude of semantic categories organized under WordNets taxonomy [4]. In particular, we utilize the table and mug categories of ShapeNet.

We used voxel models to efficiently deal with large amounts of features from point cloud data. Similar methodology can be found in Daniel Maturana and Sebastian Scherers work in [3], they propose VoxNet, an architecture by integrating a Volumetric Occupancy Grid representation with a supervised 3D Convolutional Neural Network (3D CNN).

For our model we try to learn the probabilistic latent space of tables and chairs using Wu. J, Zhang. C, Xue. T, et. al's work in [5]. All of the parameters are the same however, the generator G produces $32 \times 32 \times 32$ object, our discriminator D only accepts models with the size of $32 \times 32 \times 32$, the learning rate is .003.

METHODOLOGY

After obtaining the weakly trained models, we observe they possess some degree of original models characteristics but yet incomplete shape. The main issue is that most models have disconnected components and scattered voxel cubes in the 3D grid space due to the lack of training data. Therefore we undergo a series of morphology techniques to transform the models.

The first step was to extract the main component of the model by utilizing a depth-first search algorithm to get the largest connected component in the space.

After that, we notice there are holes in the surfaces of many models, therefore we used hole-filling algorithm to spot the holes and fix them. The algorithm used consists of invading the complementary of the shapes in input from the outer boundary of the image, using binary dilations. Holes are not connected to the boundary and are therefore not invaded. The result is the complementary subset of the invaded region.

We then take to-be-altered models from 3DShapeNet and apply a dilation algorithm to augment the models with the hole-

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single spaced.

Every submission will be assigned their own unique DOI string to be included here.

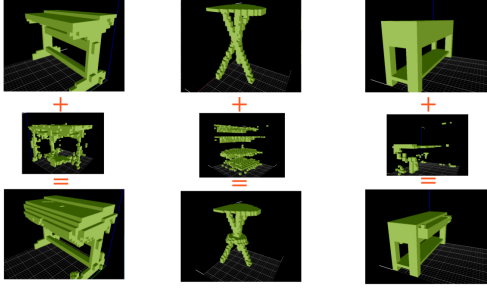


Figure 1. The top row is the original table. Second row is the GAN produced model. The third row is our augmented table.

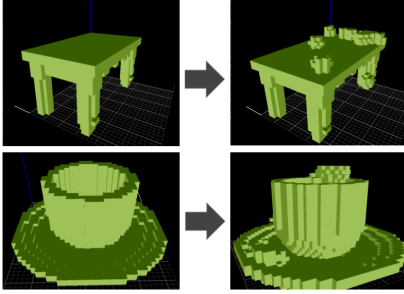


Figure 2. On the left is the original model. On the right is the poorly augmented model.

filled weakly trained models as a space constraint. We will then obtain augmented tables with certain features from the weakly trained models.

GENERATED RESULTS

Figure 1 shows examples of good results. We can see that the augmented part of model preserve some characteristics of the model. For example, some tables have drawers or some other special design added to the table. However, there are also bad results where random voxel cubes are added to the original model (see Figure 2).

We also performing image dilation using random noise to confirm that our GAN is producing useful results (see Figure 3).

EVALUATION

One of the most important factors for determining how good the results are would be to find out how visually different the generated models are from the original models. For that purpose, we ran some tests on Amazon Mechanical Turk (AMT)

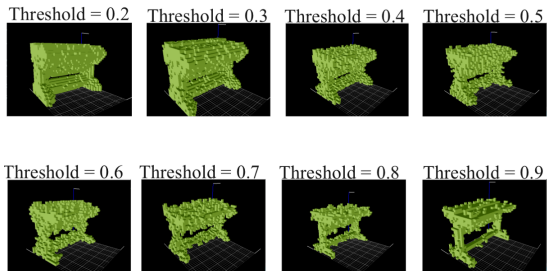


Figure 3. We vary the probability of not adding additional random noise.

where we asked users to rate the dissimilarity between the generated model and the original model on a 5 point Likert Scale. For the purpose of this experiment we chose to run our tests with tables as we believed that dissimilarities between tables would be easily discernible for the users. We recruited 30 workers for this experiment and we paid them \$0.05 per HIT. Three examples were provided to the workers for the sake of their understanding of the task. On an average, the workers rated the dissimilarity between the generated and the original model as 3.2 (fairly significant). Along with that 27/30 workers perceived the tables to be practically usable.

CONCLUSION & FUTURE WORK

To summarize we propose a method for performing data augmentation with 3D voxelized models. We utilized 2 classes from ShapeNet with a sparse number of examples. We then perform an evaluation on AMT.

In the future we plan to explore more sophisticated techniques for generating more data. For example we plan to explore morphing models directly from ShapeNet.

In addition to this we plan to share our results of training on our generated models to determine if our model learns the latent space of better. We also plan to perform a comparative analysis against training with data from sophisticated morphing techniques and augmented models using our methodology.

ACKNOWLEDGMENTS

We'd like to thank Dr. Ranjitha Kumar for supplying us with necessary resources to train our models and also for her advice and mentorship.

REFERENCES

1. Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., Xiao, J., Yi, L., and Yu, F. ShapeNet: An Information-Rich 3D Model Repository. Tech. Rep. arXiv:1512.03012 [cs.GR], Stanford University — Princeton University — Toyota Technological Institute at Chicago, 2015.
2. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2014, 2672–2680.
3. Maturana, D., and Scherer, S. Voxnet: A 3d convolutional neural network for real-time object recognition. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (Pittsburgh, PA, September 2015).
4. Miller, G. A. Wordnet: A lexical database for english. *Commun. ACM* 38, 11 (Nov. 1995), 39–41.
5. Wu, J., Zhang, C., Xue, T., Freeman, W. T., and Tenenbaum, J. B. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In

Advances in Neural Information Processing Systems
(2016), 82–90.